RESEARCH PROPOSAL

"STUDY ON STATISTICAL DATA ANALYSIS AND PREDICTION OF AGRICULTURE PRODUCTS"

A synopsis submitted to

SWAMI RAMANAND TEERTH MARATHWADA UNIVERSITY,

NANDED MAHARASHTRA

For the degree of

DOCTOR OF PHILOSOPHY

IN

STATISTICS

UNDER THE FACULTY OF SCIENCE AND TECHNOLOGY

By

Mr. Sandeep Ramesh Thorat

UNDER THE GUIDANCE OF

Dr. Vikas Shivaji Jadhav
Assistant Professor,
Department of Statistics
Sanjeevanee Mahavidyalaya, Chapoli Tq. Chakur Dist. Latur

JANUARY – 2024

"STUDY ON STATISTICAL DATA ANALYSIS AND PREDICTION OF AGRICULTURE PRODUCTS"

1. INTRODUCTION:

In India agriculture is one of the important sector of Indian economy, as per the Indian economics survey 2020-21 agriculture employed more than 50% of the Indian workforce and contributed 20.2% to the country's GDP [11]. Although the need of growing population of the country the agriculture production is much enough for the survival of Indian citizens and still remain the large exporter of many agriculture products worldwide. The diverse climatic condition in India allows Indian farmers to cultivate wide variety of crops. The price of agriculture products is the important factor for the growth of GDP. The extent of production, price instability and healthy crops are the major issues in the current agriculture scenario. Along with instability in price of agriculture products the climatic conditions, unhealthy soil conditions, irrigation facilities, the supply of quality fertilizers, cost of production are the major issues with respect to the profit of farmers. The price of agriculture production not simple to conclude. Many of the factors make impact on pricing of agriculture production such as natural calamities, demand supply conditions, National international Trade etc

Our research study aims to attempt the prediction of price of agriculture production more effectively to ensure farmers income through statistical modeling. The main attempt is to work out on seasonal and non seasonal modeling for prediction of agriculture product prices since many of the farmers can make storage of their harvested crop with their economic conditions and wait for increasing demands of products in the market so that they can maximize their profit.

Also with the market research that is by studying crop arrival pattern it is quite possible to estimate the prices of agriculture products and it becomes helpful for the farmers to plan accordingly regarding crops selection so that they can maximize their profit or minimize the loss on other side.

In this research our target is to use predictive analysis techniques to predict the prices of different agriculture products by analyzing the previous data.

2. REVIEW OF LITERATURE AND DEVELOPMENT IN THE SUBJECT:

Many researches were done in the agriculture sector intended to opt the efficient and accurate model for yield prediction in variety of crops, price prediction, weather prediction, crop disease prediction, crop selection using various statistical predictive techniques and machine learning.

Duan Yan-e,[1] proposed a system which helps us to predict the crop yield by suggesting the best crop. Jay Gholap et al., [2] this author's aim is to increase the net yield rate of the crop, based on the parameter related to the soil and atmosphere. This article helps us to predict the crop yield and suggest the best crop thereby improving the quality and prediction.

D.S.Dhakre et al., [3] studied based on the Shwartz Bayes criterion (SBC) and Akaike Information Criterion (AIC), estimated the ARIMA (2,1,1) as the estimated best model. S. Vijaybhaskar et al., [4] works to construct the model for testing the soil fertility. They also suggested the crop which has to be planted depending upon the value obtained from sensor along with the fertilizer which has to be added to the soil in order to increase the crop productivity and profit.

Pham, et al., [5] evaluated Precision agriculture (PA) techniques were used - that can help the farmers to increase productivity, reduce cost, restrict the use of chemical fertilizers and water, increase revenue, understand the seasonal trend demand. Varun, et al., [6] obtained the model that allows farmer to anticipate the best yield in terms of quality and quantity. The farmers also get the desired or the estimated price for their yield.

Vohra, et al., [7] mainly focused on proposing a decision making support model for prediction of prices in agriculture commodities . This paper also includes techniques of data mining in agriculture that will help the farmers to predict the agriculture commodity prices. Sabu , S et al., [8] studied the monthly prices of Arecaunuts in Kerala are predicted using time series and machine learning models.Wang, L. et al., [9], provide reference for the researchers by reviewing the research advances and trend of agriculture product price forecasting methods in recent years.

3. OBJECTIVES OF THE RESEARCH:

The objectives of the present study are as follows:

- To study impact of climate change on agriculture products.
- To study market arrival pattern of agriculture products.
- To develop seasonal and non seasonal predictive models for price of agricultural products.
- To develop crop selection model to maximize the yield of agriculture products.
- To study the impact of fertilizers on soil fertility.

4. METHODOLOGY:

For using the predictive analytics techniques for forecasting prices of agriculture products we require the previous year's data of agriculture products and their market prices. In general under the normal conditions prices of agriculture products follows seasonal trend. Therefore here we can opt for stochastic and time series analysis to forecast the prices of agriculture products for future seasonal and non seasonal time periods.

A) Stochastic Modeling:

Stochastic modeling develops a mathematical or financial model to derive all possible outcomes of a given problem or scenarios using random input variables. It has wide applications in various sectors like the financial market, agriculture, weather forecasting and manufacturing [15].

B) Time Series Forecasting Models for Agriculture Research:

ARIMA Model:

The Autoregressive integrated moving average – ARIMA model is a generalization of a simple autoregressive moving average –ARMA model. This model is useful for predicting future values in series based on past values [12].

An autoregressive model of order p, AR **(p)** were as follows,

$$Y_t = \delta + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} \ldots + \Phi_p Y_{t-p} + \epsilon_t$$

Where, Y_t : Value of the variable at time t.

Φ_i : Parameters of the autoregressive part of the model

ϵ_t : Error term

SARIMA Model:

The Seasonal Autoregressive integrated moving average – ARIMA model is an extension of ARIMA model that incorporates seasonality in addition to the seasonal component. SARIMA models are useful to handle the data having seasonal patterns.

HOLT-WINTERS Model:

The Holt-Winters model uses a weighted moving average to estimate local mean, trend and seasonality. For a series {Yt} with period p, the updating equations for Holt-Winters additive method are:

$$a_t = \alpha\left(Y_t - S_{t-p}\right) + (1-\alpha)(a_{t-1} + b_{t-1})$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$
$$S_t = \gamma(Y_t - a_t) + (1-\gamma)S_{t-p}$$

where a_t, b_t and S_t are the are smoothed estimates of level , trend and seasonality at time t respectively, α ,β and γ are smoothing parameters. These parameters make a distinction between the importance of recent and older observations decay exponentially.

The k-step forecasting equation for Y_{t+k} at time t is:

$$\hat{Y}_{t+k|t} = a_t + kb_t + S_{t+k-p} \text{k} \leq p$$

where a_t & b_t are the estimated level and trend at time t respectivelykb_t is the trend extrapolation; and S_{t+k-p} is the appropriate seasonal effect (i.e., in monthly series with period p=12, S_{t+k-12} corresponds to the estimated seasonal effect at the specific month)[10].

C) INVENTORY MANAGEMENT:

Inventory is any kind of resource that is stocked to satisfy the present and future needs of any organization. With respect to the agriculture sector we observe the inventory of seeds, various fertilizers, pesticides, fuel, harvested crop that is all related to farming from the cultivation phase to harvesting phase. The smart inventory management of agricultural products will be beneficial for farmers to maximize their profit [13].

D) REGRESSION MODEL:

- **Simple Linear regression model:**

  In a simple linear regression model the regressor variable depends only on a single explanatory variable.

$$Y = a + bX + e$$

  Where, Y: Regressor or dependent variable

  X: Explanatory or independent variable

  a: Intercept b: Slope & e : random error

- **Multiple Linear Regression Model:**

  In a multiple linear regression model the regressor variable depends on set of explanatory variables.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \epsilon$$

  Where,

  Y: Regressor or dependent variable

  X_i: i^{th} Explanatory or independent variable (i=1,2,…….k)

  β_0: Intercept

  β_i: Slope corresponds to i^{th} Explanatory or independent variable (i=1,2,…….k)

  ϵ : Random error.

5. IMPORTANCE OF STUDY:

The main purpose of this research is to develop the seasonal and non seasonal predictive models which will be helpful for the farmers to plan regarding the inventory of harvested crop to maximize their profit. This study will help many of those from different fields like Mathematical Science, Agriculture, Environmental Science, Market research who wishes to study Agriculture Economics.

6. PROPOSED WORK PLAN:

Sr. No.	Period	Plan of the Work
1.	First Year	• Study of Background for the research topic. • To Collect and analyze some more research articles on the concerned subject. • Try to publish a review paper on concerned topic.
2.	Second Year	• Data collection and preprocessing. • To develop mathematical and statistical models. • Data analysis and interpretation of results. • Publishing research articles in national and international journals.
3.	Third Year	• Interpretation of results. • Writing of final thesis and submission to University.

7. REFRENCES :

[1] Duan Yan-e,(2011) "Design of Intelligent Agriculture Management Information System Based on IOT‖",IEEE,4th,.*Fourth International reference on Intelligent Computation Technology and Automation.*

[2] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, (May 2012) "Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction", *IJCSI*, Vol. 9, Issue 3, No 3, ISSN: 1694-0814.

[3] D.S.Dhakre and D. Bhattacharya, (May 2014) " Price Behaviour of Potato in Agra Market-A Statistical Analysis", *Indian Res. J. Ext. Edu. 14 (2)* .

[4] P.S.Vijaybaskar, R. Sreemathi and E. Keertanaa,(2017) "Crop Prediction using Predictive analytics" *International Conference on Computation of Power,Energy information and Communication (ICCPEIC)*. .

[5] Pham, Xuan Stack , Martin, (2018) " How data analytics is transforming agriculture." Business *Horizons ,Elsevier* , vol. 61(1), pages 125-133.

[6] Varun, R., N, N., HP, S., A, S., & M, M. (2019). " Agriculture Commodity Price Forecasting using Ml Techniques". *International Journal of Innovative Technology and Exploring Engineering*, 9(2S), 729–732.

[7] Vohra, A., Pandey, N., & Khatri, S. K.(2019) "Decision Making Support System for Prediction of Prices in Agricultural Commodity". Proceedings - 2019 *Amity International Conference on Artificial Intelligence, AICAI* 2019, 345–348.

[8] Sabu, K. M., & Kumar, T. K. M.(2020) Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala. *Procedia Computer Science*, 171, 699–708.

[9] Wang, L., Feng, J., Sui, X., Chu, X., & Mu, W. "Agricultural product price forecasting methods: research advances and trend". *British Food Journal*.

[10] Gustavo R.A.B, Anaily R.V., Andres L Q, & Maria E.R.P., (2021), " Comparison between SARIMA and Holt-Winters models for forecasting monthly streamflow in western region of Cuba"3,671.

[11] *en.wikipedia.org/wiki/Agriculture_in_India* assessed on 6 Jan. 2024

[12] D.C. Montgomerry, C.L Gennings, MKulachi,(2010),Tntroduction to Time Series Analysis and Forecasting" *A John Wiley& Sons, INC., Publication.*

[13] J.K. Sharma, (2007) ,"Operations Research Theory and Applications" 6th edition, *Trinity publication.*

[14] D. C. Montgomery , E.A. Peck, G.G. Vining,(2012), " Introduction to Linear Regression Analysis", *5th edition, Wiley Publication.*

[15] S.M. Ross , (2008) ," Stochastic Processes" , *Wiley Student Edition.*

Mr. Sandeep Ramedh Thorat
(Research Student)

Dr. Vikas Shivaji Jadhav
(Research Guide)